

THE COMPLEMENTARITY BETWEEN CITIES AND SKILLS*

Edward L. Glaeser

Department of Economics, Harvard University, 315A Littauer Center, Cambridge, MA 02138. Also affiliated with NBER. Email: eglaeser@harvard.edu.

Matthew G. Resseger

Department of Economics, Harvard University, Littauer Center, Cambridge, MA 02138. Email: resseger@fas.harvard.edu.

ABSTRACT. There is a strong connection between per-worker productivity and metropolitan area population, which is commonly interpreted as evidence for the existence of agglomeration economies. This correlation is particularly strong in cities with higher levels of skill and virtually nonexistent in less skilled metropolitan areas. This fact is particularly compatible with the view that urban density is important because proximity spreads knowledge, which either makes workers more skilled or entrepreneurs more productive. Bigger cities certainly attract more skilled workers, and there is some evidence suggesting that human capital accumulates more quickly in urban areas.

1. INTRODUCTION

The connection between area size and per worker productivity and income is a core fact at the center of urban economics (Glaeser, 2008). The connection between urban density and earnings is understood to be a primary reason that cities exist. Understanding the connection between city size and productivity is a core task for students of agglomeration.

This paper notes that the connection between city size and productivity does not hold for less skilled metropolitan areas in the United States today. In the least well-educated third of metropolitan areas, there is virtually no connection between city size and productivity or income. In the most well-educated third of metropolitan areas, area population can explain 45 percent of the variation in per-worker productivity.

Why does productivity increase with area population for skilled places, but not for unskilled places? One hypothesis is that the connection between productivity and area size reflects a tendency of more skilled people to locate in big cities. However, even in the more skilled places, controlling for area-level skills can only explain a quarter of the measured agglomeration effect. If unobserved skills were explaining the correlation, then we would expect

*Both authors thank the Taubman Center for State and Local Government for research support. Kristina Tobio provided excellent research assistance; Erin Dea provided excellent editorial aid. Yuko Aoyama and Gilles Duranton provided helpful comments.

Received: April 2009; revised: May 2009; accepted: May 2009.

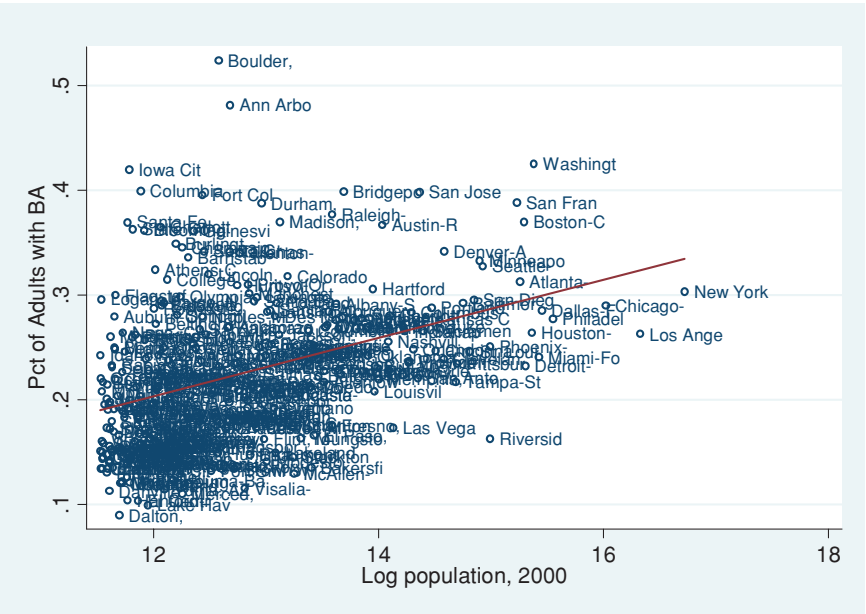
real wages to rise with city population, which they do, but that effect seems to explain only 30 percent of the connection between city size and income or productivity.

We divide the theories of agglomeration into two broad categories: those that emphasize the spread of knowledge in cities and those that do not. Among the latter group is the view that cities are more productive because of advantages unrelated to agglomeration, such as access to ports or harbors or good government, and the possibility that capital is more abundant in big cities. Nonknowledge-based theories also include standard agglomeration models, where urban proximity reduces transport costs. In Section 3, we address these theories. While there is little evidence that directly supports these hypotheses, there is little evidence with which to reject them either.

In Section 4, we turn to two core knowledge-based theories of urban agglomeration, which can both readily explain why the productivity-city size connection is so much stronger in higher human capital metropolitan areas. The first hypothesis, which comes from Marshall's statement (1890) that in agglomerations the "mysteries of the trade" are "in the air," is that density makes it easier for workers to learn from each other. The second hypothesis is that high levels of human capital and city size interact to push out the frontier of knowledge and the level of productivity. While these two hypotheses predict similar things about the links between productivity, human capital and city size, two natural versions of the theories have different implications for wage growth in skilled cities. The learning interpretation suggests that age-earnings profiles should be steeper in big, skilled areas, because workers are learning more rapidly. One version of the innovation interpretation implies that age-earnings profiles in such places are flatter, because technological change is proceeding rapidly and making the skills of older people obsolete. This implication requires the added assumption that technological change causes some skills to become out of date.

As in Glaeser and Mare (2001), we find some evidence supporting the view that workers learn more quickly in metropolitan areas. We also find that this learning effect is stronger in more skilled areas. However, we do not find that age-earnings profiles are steeper in bigger metropolitan areas, and the interaction between area size, area skills, and experience is insignificant. While these findings are quite compatible with the view that cities and skills are complements, they do not clearly indicate whether this complementarity works through learning, innovation, both or neither.

The natural implication of the view that cities and human capital are complements is that cities will become more, not less, important if humanity continues acquiring knowledge. The importance of connecting in dense urban areas will only increase if knowledge becomes more important, at least as long as technological shifts do not eliminate the urban edge in transferring information.



Notes: (i) Units of observation are MSAs under the 2006 definitions with populations above 100,000. Statistics are from the Census.
(ii) The regression line is $Share\ with\ BAs = 0.028\ [0.003] * Log\ population - 0.13\ [0.044]$. $R^2 = 0.16$ and $N = 335$.

FIGURE 2: Area Skill and Area Size.

disproportionately increase the productivity of more skilled workers. These two theories can be distinguished; if this connection reflected a demand for amenities it would mean that cities are skilled because of abundant labor supply, and we should expect to see lower wages for skilled workers in big cities (Glaeser, 2008). A naïve attempt to control for the share of adults with college degrees at the metropolitan area level yields the following regression

$$\begin{aligned} \text{Log(Output per Worker)} &= 9.49 + 0.098 * \text{Log(Population)} \\ &\quad (0.11) \qquad (0.010) \\ (1) \qquad \qquad \qquad &+ 1.18 * \text{Share with BAs.} \\ &\qquad \qquad \qquad (0.14) \end{aligned}$$

Output per worker continues to be gross metropolitan product divided by the size of the labor force. Standard errors are in parentheses. The R^2 is 0.47 and there are 335 observations. The coefficient on log of population declines slightly, from 0.13 to 0.098, roughly a 25 percent decline. Just controlling for human capital eliminates about one-quarter of the connection between area population and output per worker.

But it appears that the effects of human capital and city size are not independent. When we interact the two variables, we estimate

$$\begin{aligned} \text{Log(Output per Worker)} = & 0.08^* \text{Log(Population)} + 1.26^* \text{Share with BAs} \\ & (0.01) \qquad (0.14) \\ (2) \qquad \qquad \qquad & + 0.51^* \text{BA}^* \text{Pop.} \\ & (0.13) \end{aligned}$$

The term $\text{BA}^* \text{Pop}$ refers to the product of log of area population (demeaned) and share with college degrees (also demeaned). An intercept was included in the estimation but is not reported for space reasons. The R^2 is now 0.49. The demeaning of the variables means that both raw coefficients can be interpreted as the impact of the variable, when the other variable has taken on its mean level. The interaction means that when the share with college degrees is at its minimum observed value of 0.09 (which would be -0.13 relative to the mean), the estimated coefficient on population is just 0.01, whereas for the maximum value of 0.52, the estimated effect is 0.23.

If we instead run this regression with the logarithm of per capita income, we estimate coefficients of 0.026 on the log of population, 1.43 on share of the population with college degrees and 0.42 on the interaction. In this income specification, the t -statistic on the interaction is 4.5. If we use log of median family income as the dependent variable, the estimated coefficients are 0.019, 1.55 and 0.36 on the three variables. The t -statistic on the interaction remains over 4.

Our independent variables are certainly endogenous, and we have no perfect source of exogenous variation that solves this problem. However, similar results appear if we use variables from 1940 (population, share with college degrees, and the interaction) instead of contemporaneous variables to explain current gross metropolitan product. In that case, we estimate

$$\begin{aligned} \text{Log(Output per Worker)} = & 0.07^* \text{Log(Population)} + 5.04^* \text{Share with BAs} \\ & (0.01) \qquad (0.68) \\ (2') \qquad \qquad \qquad & + 2.47^* \text{BA}^* \text{Pop.} \\ & (0.60) \end{aligned}$$

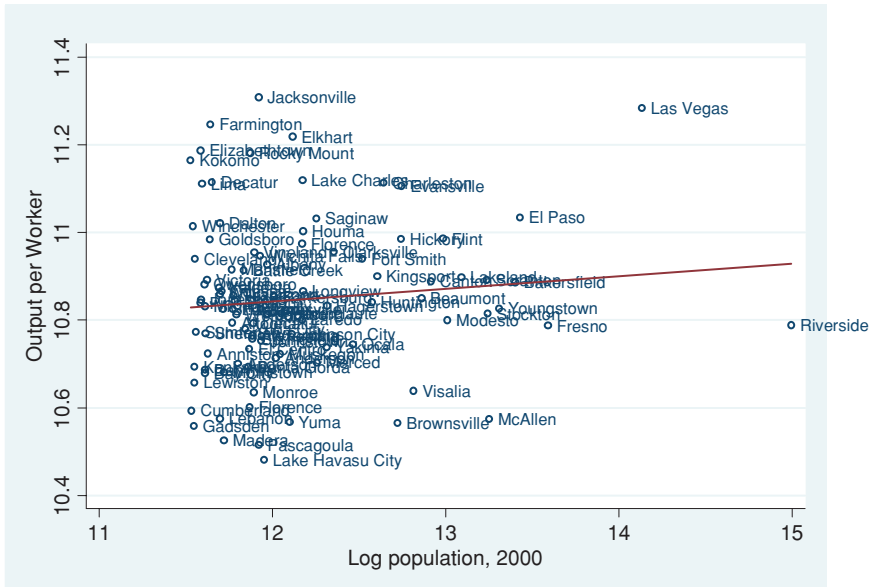
In this case, there are 334 observations and the R^2 is 0.34. The high coefficient on the lagged share of the population with college degrees reflects, in part, the tendency of skilled places to become more skilled over time, as discussed in Berry and Glaeser (2005).

In individual-level regressions, which control for individual-level human capital and experience, our results weaken significantly. The first regression of Table 1 shows the 0.041 coefficient when individual yearly log earnings are regressed on metropolitan area size (also found in Glaeser and Gottlieb, 2008). This coefficient is less than one-half of the baseline coefficient estimated in the aggregate gross metropolitan product regression. Controlling for the share of the population with college degrees pushes the coefficient down further to

TABLE 1: Log Annual Income on City Population and Skill Levels Interacted with Experience

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log population	0.041 (0.011)	0.028 (0.011)	0.022 (0.012)	0.038 (0.006)	0.034 (0.016)	0.044 (0.019)	0.021 (0.018)
Pct with BA		0.639 (0.144)	0.411 (0.122)	0.208 (0.086)	0.415 (0.123)	-0.11 (0.318)	-0.895 (0.287)
LogPop*PctBA interaction			0.196 (0.113)	0.413 (0.087)	0.193 (0.113)	0.193 (0.113)	0.885 (0.229)
LogPop*LogExp					-0.004 (0.004)	-0.007 (0.005)	0 (0.004)
PctBA*LogExp					0.18 (0.095)	0.18 (0.095)	0.448 (0.086)
LogPop*PctBA*LogExp							-0.236 (0.057)
Log experience	0.25 (0.004)	0.252 (0.004)	0.252 (0.004)	0.251 (0.005)	0.258 (0.006)	0.256 (0.005)	0.254 (0.005)
Education Dummies:							
0-9 years	-0.59 (0.010)	-0.587 (0.009)	-0.586 (0.009)	-0.584 (0.011)	-0.586 (0.009)	-0.585 (0.009)	-0.585 (0.009)
10-11 years	-0.33 (0.005)	-0.327 (0.005)	-0.327 (0.006)	-0.319 (0.006)	-0.327 (0.006)	-0.327 (0.006)	-0.326 (0.006)
13-15 years	0.207 (0.004)	0.204 (0.004)	0.204 (0.004)	0.179 (0.004)	0.204 (0.004)	0.204 (0.004)	0.204 (0.004)
16 years	0.575 (0.008)	0.565 (0.008)	0.566 (0.008)	0.516 (0.006)	0.566 (0.008)	0.566 (0.008)	0.566 (0.008)
17+ years	0.788 (0.011)	0.774 (0.010)	0.774 (0.010)	0.717 (0.008)	0.774 (0.010)	0.774 (0.010)	0.775 (0.010)
Constant	9.409 (0.015)	9.406 (0.015)	9.406 (0.015)	9.41 (0.018)	9.388 (0.019)	9.394 (0.018)	9.401 (0.017)
Observations	2,102,175	2,102,175	2,102,175	2,102,175	2,102,175	2,102,175	2,102,175
R ²	0.16	0.17	0.17	0.15	0.17	0.17	0.17

Note: Robust standard errors in parenthesis.



Notes: (i) Units of observation are MSAs under the 2006 definitions with populations above 100,000 and where the share of adults with college degrees is less than 17.65 percent. Labor force and population is from the Census, as described in the Appendix. Gross Metropolitan Product is from the Bureau of Economic Analysis.

(ii) The regression line is $\text{Log GMP per capita} = 0.028 [0.028] * \text{Log population} + 10.50 [0.34]$. $R^2 = 0.01$ and $N = 100$.

FIGURE 3: Productivity and Area Size in Less Skilled MSAs.

0.028. In the third regression, we show that the interaction between population and the share with college degrees is positive, although significant only at the 10 percent level.

The individual-level results are qualitatively similar to those above although weaker in magnitude. The differences between the individual and aggregate regressions reflect primarily the fact that the aggregate results are weakest for the largest metropolitan areas, which are weighted heavily in these individual-level regressions. Regression (4) repeats regression (3) weighting by the inverse of Metropolitan Statistical Area (MSA) population (so smaller metropolitan areas get more weight). In this case, the results look similar to the aggregate results.

Figures 3 and 4 show the interaction between output per worker and metropolitan area population graphically. Figure 3 shows this relationship in the 100 least well-educated metropolitan areas with populations over 100,000. Figure 4 shows the relationship between metropolitan area population and output per worker in the 100 most well-educated areas with populations over 100,000. Among less well-educated places, there is essentially no

population is particularly strong in such places.² We will address the theory that cities enhance skill acquisition later. Here, we just discuss the possibility that the urban wage premium reflects preexisting skills. After all, as Bacolod, Blum, and Strange (2009) emphasize in this volume, skills are far more than years of education. Glaeser and Mare (2001) do a fair amount of work showing that the urban wage premium (as opposed to the more continuous correlation between city size and productivity or earnings) survives a large number of measures of individual human capital, such as test scores and instrumental variables approaches that use parental state of birth characteristics.

One of their pieces of evidence supporting the view that omitted premarket human capital variables are not higher in cities is that real wages, that is, wages controlling for local price levels, do not rise significantly in urban areas. If people in cities had higher levels of innate human capital, then they should be earning higher real wages as well as higher nominal wages. After all, they are more skilled. Of course, estimated real wages would need to be adjusted for local amenities, and amenities may be either higher or lower in large urban areas.³ Glaeser and Mare find little connection between city size and real wages in their sample of cities. In our considerably larger sample, we also find little connection between the log of median family income, divided by the American Chamber of Commerce Research Association local price index, and city population, at least once we control for the share of the population with college degrees.

However, this result is not true in the more skilled cities where agglomeration elasticities are strongest. For example, if we look only at those areas where the share of population with college degrees is greater than 25.025 percent (the same cutoff used to establish the top 100 skilled cities above), we find that

$$\begin{aligned} \text{Log(Real Family Income)} = & 9.07 + 0.025 * \text{Log(Population)} \\ & \quad (0.10) \quad \quad (0.006) \\ (3) \quad & + 1.00 * \text{Share with BAs.} \\ & \quad \quad (0.15) \end{aligned}$$

There are 100 observations and the R^2 is 0.27. All data come from the Census except for the price indices used to turn nominal into real income, which come from the American Chamber of Commerce Research Association.⁴

Real incomes rise significantly with skills, which is compatible with the view that more skilled people are more productive. While real incomes do not rise with city size, across the entire population, in these skilled areas, there is a

²If skills were learned in big cities, then more human capital in big cities would lead to more learning in the model of Glaeser (1999). If skills were preexisting, then it would be possible that omitted aspects of human capital were more important at the high end of the skill distribution which is overrepresented in skilled places.

³Glaeser (2008, Chapter 3) presents a lengthy discussion of the spatial equilibrium, Rosen–Roback model, which underlies this logic.

⁴A better procedure would be to use individual-level data and individual-level price controls as in Moretti (2008).

positive connection. This connection can be interpreted as either implying that there is a greater level of unobserved human capital in these areas or that these bigger cities are particularly unpleasant and higher wages are compensation for negative amenities. However, controlling for some obvious amenities, such as temperature, does little to change this result, and we have trouble believing that there are more negative amenities in big skilled cities than in big unskilled cities.⁵

If the coefficient on city size is treated as a measure of the extent to which unobserved skills rise with city size in this skilled city subsample, this would mean that about 30 percent of the urban productivity coefficient could be explained by human capital (0.025/0.08).⁶ Since observed human capital is uncorrelated with city size in this subsample, this is a plausible measure of the extent to which human capital explains the city size effect in these cities. In the larger sample that includes skilled and unskilled cities, bigger cities do have higher observed levels of human capital, and controlling for skills can explain about one-quarter of the connection between city size and productivity, but there is little sign that unobserved human capital is higher in bigger metropolitan areas in that larger sample of cities. In either case, human capital appears to explain at most 30 percent of the city size effect, leaving at least 70 percent to be explained.⁷ Understanding why the city size effect is larger in skilled places seems particularly pressing.

3. URBAN PRODUCTIVITY FRAMEWORK

We now use a standard production function to consider alternative interpretations of our agglomeration results.⁸ In a standard production function, output per worker can be written as $PAF(K, hL)/L$, where P is the price of the good, A is the level of productivity, K is the level of capital, and hL reflects the amount of effective labor, with h as human capital and L as the number of workers. If the production function is homogenous of degree one, which is necessary for a zero-profit equilibrium, then output per worker can be rewritten as $PAF(k, h)$, where k reflects physical capital per worker and h reflects human capital. If the production function is Cobb-Douglas, with parameter β on labor, then differentiating this quantity with respect to any exogenous variable " Z ", such as city population, yields the following decomposition:

⁵For example, the problem of urban crime is particularly prevalent in less skilled metropolitan areas.

⁶Note that this real-wage method would only get at exogenous unobserved skills. If cities created unobserved skills endogenously, then in a spatial equilibrium, workers should end up paying for those skills with higher costs of living (Glaeser and Mare, 2001).

⁷Combes, Duranton, and Gobillon (2008) find that unobserved skills can explain up to one-half of the connection between agglomeration size and wages in France. The discrepancy between their results and our results here might reflect differences between the United States and France or their use of individual fixed effects to control for unobserved skills.

⁸Puga (2009), in this volume, provides a more thorough discussion of the different sources of agglomeration economies.

$$(4) \quad \frac{\partial \text{Log}(\text{Output Per Worker})}{\partial \text{Log}(Z)} = \frac{\partial \text{Log}(P)}{\partial \text{Log}(Z)} + \frac{\partial \text{Log}(A)}{\partial \text{Log}(Z)} + (1 - \beta) \frac{\partial \text{Log}(k)}{\partial \text{Log}(Z)} + \beta \frac{\partial \text{Log}(h)}{\partial \text{Log}(Z)}.$$

Wages per worker equal the wage per effective unit of human capital times the amount of human capital per worker. In a standard Cobb-Douglas formulation, wages per worker equal β times output per worker.

To close the model, capital and labor should also be endogenized. If workers are to be indifferent across locations, which is a necessary condition for the existence of a spatial equilibrium (Glaeser, 2008), then high costs of living must offset high wages. But that fact does not change the fact that high wages must also be offset by something making firms more productive, and our focus is on this latter relationship.

The connection between output per worker and city size could represent an increase in prices, productivity, capital per worker, or human capital per worker, in big cities. The relative importance of the different forces will surely differ across industries. Barbers will have a higher output per worker in bigger cities, but much of that difference will reflect higher prices, not capital per worker, or even human capital. Conversely, the prices of traded manufactured goods are more or less constant over space, and any variation in output per worker in that industry is likely to reflect productivity or capital, either physical or human.

We will divide up these theories into two sets of hypotheses. One set of theories emphasizes greater knowledge in cities, which could mean higher levels of “ h ” or a higher level of “ A ” brought on by the urban exchange of ideas. We will address that set of theories in the next section. The other set of theories focuses on other causes of urban productivity, which include innate urban advantages, such as access to waterways or good government, higher levels of capital per worker, and non-knowledge-based agglomeration economies.

Conceptually, it would be quite possible for the strong connection between city size and productivity to reflect omitted characteristics of a location that both enhance productivity and attract workers. In the 19th century, it seems undebatable that the waterways of New York and Chicago made these places economically successful and attracted people to them (Glaeser, 2005). Yet few urbanists believe that locational advantages have much direct impact on productivity today.⁹ Cities long ago gave up on those industries that were tied to their local geography. Today, cities are more likely to specialize in business services (Kolko and Neumark, 2008), and it is hard to see how those services get an edge from a harbor or a coal mine. Natural advantages seem to explain only 25 percent of the concentration of manufacturing industries (Ellison and Glaeser, 1999).

⁹Combes et al. (2008) use historical sources of innate advantage as instruments for current population density, which requires that these variables be orthogonal to current productivity.

We are less sure that natural advantage is irrelevant in explaining the connection between city size and productivity, but it seems unlikely that any natural advantages can explain why that connection is stronger in more skilled cities. After all, many of these natural advantages would seem to have their largest impact on less skilled industries. Indeed, that is exactly what a cursory examination of the data reveals. Variables like proximity to the great lakes or harbors positively impact productivity in less skilled places, but have no impact in more skilled areas. For example, the correlation between per capita income and miles from the nearest body of water is -0.33 for less educated cities and -0.03 for more educated cities. If this result holds more generally, and innate advantage matters more for less skilled workers, then the fact that city size increases productivity more for places with more skills is evidence against the importance of such natural advantages.

One way in which natural advantage might matter today is that past historical natural advantages might have led to more investment in physical capital. Typically, physical capital is treated as endogenous and for that reason, not really a plausible determinant of agglomeration economies. For example, in the model sketched above, if purchased by producers at a cost " r ", which might differ across space, then a Cobb-Douglas relationship would imply that:

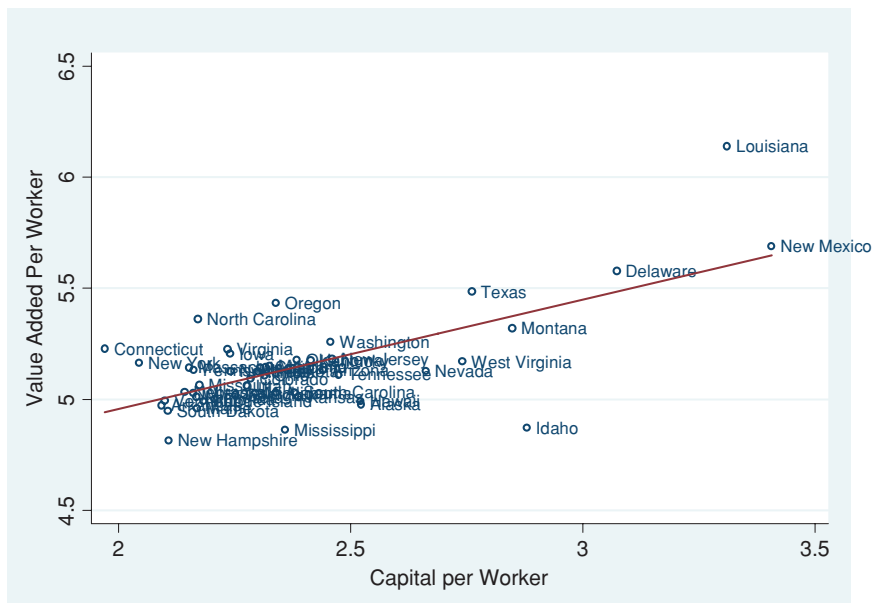
$$(4') \quad \frac{\partial \text{Log}(\text{Output Per Worker})}{\partial \text{Log}(Z)} = \frac{1}{\beta} \left(\frac{\partial \text{Log}(P)}{\partial \text{Log}(Z)} + \frac{\partial \text{Log}(A)}{\partial \text{Log}(Z)} \right) - \frac{1 - \beta}{\beta} \frac{\partial \text{Log}(r)}{\partial \text{Log}(Z)} + \frac{\partial \text{Log}(h)}{\partial \text{Log}(Z)}.$$

If physical capital is endogenously determined, then it can only increase the connection between city size and output per worker if capital is cheaper in big, dense cities. Typically, evidence on real estate costs would suggest that capital is, if anything more expensive in big cities, which reflects the greater scarcity of land.

However, if big cities have long invested in durable physical capital, then that capital might remain and might increase productivity today. Certainly, casual observation of cities such as New York, London, and Paris suggests that they have advantages that come from centuries of public and private investment in physical capital. Is there any evidence to support this view?

Unfortunately, there is little good measurement of physical capital at the metropolitan area level. A few heroic social scientists, such as Munnell (1991) and Garofalo and Yamarik (2002) have created state-level estimates of the capital stock, but these estimates have been based on apportioning the national capital stock to states on the basis of the types of industries in those states.¹⁰ At the state level, for manufacturing industries, the Census of Manufacturers

¹⁰A similar procedure could be used at the metropolitan area level, but we doubt that it would be seen as particularly compelling to suggest that New York City's capital stock is the same as the nation, except for its mix of major industries.



Notes: (i) Measures of value added and capital per worker are taken from the Census of Manufacturers as described in the Appendix.

(ii) The regression line is $\text{Log Value Added per Worker} = 0.4924 [0.079] * \text{Log Capital per Worker} + 3.97 [0.191]$.

$R^2 = 0.45$ and $N = 49$.

FIGURE 5: Value-Added and Physical Capital.

provides an estimate of expenditures on capital. These numbers are problematic in two ways: they represent an estimate of the flow of investment not the stock of capital and they address only manufacturing.

While there is certainly a robust relationship between capital expenditures and value added per worker, shown in Figure 5, controlling for capital expenditures only increases the relationship between state-level density and value added per worker or income. Table 2 shows the relationship between the Ciccone and Hall (1996) index of state-level density and two measures of output: value added per worker and hourly wage for production workers for states with more than 50,000 manufacturing workers. Columns (2) and (4) show the increased connection between output and density when a control for capital expenditures is added. The raw correlation between capital expenditures and density is negative. These results should not lead us to think that the capital stock explanation for urban productivity is disproved, but rather that this sliver of available evidence does not support that hypothesis.

In columns (3) and (6), we include controls for years of schooling taken, for comparability reasons, also from Ciccone and Hall (1996). Schooling has a

TABLE 2: State-Level Density and Output

	Log Value Added			Log Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
State-level density	0.585 [0.304]	0.787 [0.314]	0.559 [0.337]	0.458 [0.178]	0.551 [0.188]	0.342 [0.192]
Log capital per worker		0.244 [0.132]			0.112 [0.0789]	
Years of schooling			0.0177 [0.0911]			0.0775 [0.0518]
Constant	4.339 [0.397]	3.516 [0.588]	4.137 [1.111]	2.297 [0.233]	1.919 [0.351]	1.416 [0.632]
Observations	37	37	37	37	37	37
R^2	0.096	0.178	0.097	0.158	0.205	0.21

Notes: (i) Standard errors in parenthesis.

(ii) State-level density and years of schooling from Ciconne and Hall (1996).

(iii) Log value added, log wage, and log capital per worker from 2006 Annual Survey of Manufactures at factfinder.census.gov.

tiny and insignificant impact on value added per worker, and a larger but still insignificant effect on the hourly wage. Controlling for schooling reduces the coefficient on density in the wage regression, but not the value-added regression, because education seems to influence wages more than value added.

Still, this evidence only informs us about current expenditures, not the stock of accumulated urban capital. We know of no good measures of such historical investment, but we can at least ask whether historical development eliminates either the current link between population and productivity or the interaction between that variable and the share of the population with college degrees. Including the logarithm of population in 1900 as a control in regression (2) yields

$$\begin{aligned}
 \text{Log(Output/Worker)} = & 0.72^* \text{Log(Pop}_{.2000}) + 1.23^* \text{Share with BAs} \\
 & \quad (0.01) \quad (0.14) \\
 (2'') \quad & + 0.53^* \text{BA}^* \text{Pop} + 0.15^* \text{Log(Pop}_{.1900}). \\
 & \quad (0.13) \quad (0.008)
 \end{aligned}$$

There are many problems with this regression, including the fact that population growth between 1900 and today is hardly random, but its results give little hope to the view that historical investment in capital stock explains either the basic agglomeration effect or the interaction between education and population. Neither coefficient is substantially changed from (2). We have also experimented using geographic instruments, like proximity to the Great Lakes or rivers navigable in 1900, which do predict population in that year, but instrumental variables regressions show little change relative to the ordinary least squares regressions. As such, we find little evidence to support the view

that greater capital in cities explains much.¹¹ Equations (4) and (4') leave us with two alternative views about the connection between productivity and area size. In principle, the equations suggest that either higher prices or standard agglomeration effects, coming from reductions in transport costs, could explain the productivity-area size link. We believe that these two views can be taken together, since in many cases, higher prices are directly reflecting agglomeration economies. For example in Krugman (1991), concentrated firms are able to get more for their goods because other firms are located in the same area. Prices will actually be lower in the core as well, because transport costs are saved, and lower costs of intermediate goods can also increase productivity.

There is a long and distinguished literature on agglomeration economies, and there is little doubt that many forms of such economies exist. Such traditional agglomeration effects are compatible with the absence of agglomeration economies in low human capital cities only if there is some reason why the industries in those cities do not benefit from proximity, while industries in high-human capital cities do. Yet controlling for the industrial characteristics of the metropolitan area, and for interactions between these variables and area population, has little impact on the robust interaction between population and skill levels. It is not clear if all of these theories can explain the interaction between city size and human capital, but at least some of them can. For example, if agglomeration economies came from the reduction of transaction costs in business services, and if those costs took the form of lost time, then the value of reducing these time costs would be higher in place with higher levels of human capital. If these standard agglomeration economies explain the city size-productivity link, then hopefully future work will help us to understand why these effects are stronger in more educated places.

4. THE LINK BETWEEN HUMAN CAPITAL AND AGGLOMERATION ECONOMIES

While standard agglomeration theories do not automatically predict the interaction between urban size and area education, theories that emphasize the spread of knowledge in urban areas do. If cities facilitate the spread of information, then this advantage will be more important when the people living in those cities have higher levels of human capital. This suggests that there are two, in some senses quite similar, hypotheses that can explain the overall connection between productivity and agglomeration and why the agglomeration effects are so much stronger in skilled places. One view is that workers acquire more skills in big, skilled areas (Glaeser, 1999; Peri, 2002). The second view is that the Solow residual is higher in such places because of the speedy

¹¹These estimates primarily focus on private capital. Many forms of public capital, such as roads, are disproportionately present in less dense areas (Duranton and Turner, 2008), so we suspect that controlling for public capital would do little to explain the connection between density and productivity.

spread of ideas. According to the first of these theories, the workers on Wall Street benefit from the ability to learn more quickly from each other. According to the second view, their firms' leaders are better able to acquire ideas in these areas.

While this latter hypothesis has been taken seriously since Lucas (1988), we know of little direct evidence testing this view.¹² There has been more work on the connection between worker human capital accumulation and urban density. The two views differ in their predictions about the age-earnings profile in cities. The worker-learning hypothesis suggests that age-earnings profiles should be steeper in skilled, dense areas where workers learn from each other. The innovation hypothesis can mean that skills depreciate more quickly in such places, which would make the age-earnings profile flatter. We test to distinguish these two hypotheses here.

Evidence on Worker Learning in Cities

Glaeser and Mare (2001) examine the urban wage premium in models with worker-fixed effects. They find that only a modest fraction of the urban wage premium is earned by workers when they come to urban areas. Similarly, the urban wage premium is not lost by workers when they leave big cities. Instead, workers who came to cities experience somewhat faster wage growth. This evidence seems to point against a generalized urban-productivity effect toward a wage-growth effect, which could be interpreted as faster learning in cities. This wage growth may also be associated with easier job hopping in cities, where workers increase wages and productivity as they move from firm to firm (Freedman, 2008). The connection between wage gains and job mobility may reflect better matching in cities, or the gradual accumulation of human capital that is acquired when individuals work for different employers.

Since human capital accumulation is typically inferred by looking at age-earnings profiles, it is particularly natural to test the hypothesis that cities increase the rate of human capital accumulation by looking at whether wage growth is faster over the life-cycle in metropolitan areas. Table 3 shows the basic pattern of wage growth in urban areas. The dependent variable is the log of hourly wage, and data come from the 2000 Census. The first column shows the basic pattern of wage growth over the life-cycle for males between the ages of 25 and 65 (to avoid retirement issues and working part time). Experience is defined as age minus years of education minus six.

The first column shows that the majority of earnings growth occurs over the first 15 years. Relative to workers with between 0 and 5 years of experience, workers with between 6 and 10 years of experience earn 0.194 log points higher wages and workers with between 11 and 15 years of experience gain 0.335 log

¹²Relatively little work has been done using microdata to assess whether firm productivity rises with time in a dense, or well-educated city. Breau and Rigby (2008) look at such a learning model, but focus on learning through exporting.

TABLE 3: Log Hourly Wage on the Interactions of Metropolitan Residence and Human Capital Variables

	Basic Human Capital Regression (1)	Metro Areas vs. NonMetro (2)	Highly Educated Metro Area vs. NonMetro (3)	Low Educated Metro Area vs. NonMetro (4)	Cols. 3 & 4 differ at 1% level (5)	Highly Populated vs. Less Populated MSAs (6)	High Pop vs. Less Pop: Highly Educated MSAs (7)
In metro area		0.034 (0.007)					
In high educ MSA			0.069 (0.007)		Yes		
In low educ MSA				0.015 (0.011)			
In high pop MSA						0.081 (0.006)	0.088 (0.008)
Experience Dummies:							
0–5 years	(Omitted)	(Omitted)	(Omitted)	(Omitted)		(Omitted)	(Omitted)
6–10 years	0.194 (0.003)	0.028 (0.007)	0.035 (0.007)	0.004 (0.012)	Yes	0.002 (0.006)	–0.005 (0.008)
11–15 years	0.335 (0.003)	0.06 (0.007)	0.075 (0.007)	0.017 (0.012)	Yes	0.006 (0.006)	–0.009 (0.008)
16–20 years	0.423 (0.003)	0.074 (0.007)	0.093 (0.007)	0.027 (0.012)	Yes	0.01 (0.006)	–0.014 (0.008)
21–25 years	0.466 (0.003)	0.074 (0.007)	0.089 (0.007)	0.044 (0.012)	Yes	0.007 (0.006)	–0.009 (0.008)
26–30 years	0.493 (0.003)	0.067 (0.007)	0.077 (0.007)	0.043 (0.012)	Yes	0 (0.006)	–0.014 (0.008)
31–35 years	0.523 (0.003)	0.075 (0.007)	0.084 (0.008)	0.053 (0.012)	Yes	–0.001 (0.006)	–0.019 (0.008)
36–40 years	0.535 (0.003)	0.067 (0.008)	0.076 (0.008)	0.046 (0.013)	No	–0.001 (0.007)	–0.013 (0.009)
41+ years	0.515 (0.003)	0.079 (0.008)	0.09 (0.009)	0.051 (0.013)	Yes	0 (0.007)	0 (0.010)

Continued

TABLE 3: Continued

	Basic Human Capital Regression (1)	Metro Areas vs. NonMetro (2)	Highly Educated Metro Area vs. NonMetro (3)	Low Educated Metro Area vs. NonMetro (4)	Cols. 3 & 4 differ at 1% level (5)	Highly Populated vs. Less Populated MSAs (6)	High Pop vs. Less Pop: Highly Educated MSAs (7)
Education Dummies:							
0–9 years	–0.297 (0.002)	–0.047 (0.005)	–0.055 (0.005)	–0.06 (0.007)	No	–0.036 (0.005)	–0.026 (0.007)
10–11 years	–0.152 (0.002)	–0.007 (0.004)	–0.006 (0.004)	–0.015 (0.006)	No	–0.011 (0.004)	–0.02 (0.006)
13–15 years	0.108 (0.001)	0.025 (0.003)	0.021 (0.003)	0.026 (0.004)	No	0.015 (0.003)	0.01 (0.004)
16 years	0.304 (0.002)	0.093 (0.004)	0.093 (0.004)	0.032 (0.007)	Yes	0.015 (0.003)	0.004 (0.005)
17+ years	0.407 (0.002)	0.099 (0.006)	0.095 (0.006)	0.062 (0.010)	Yes	0.016 (0.005)	0.008 (0.006)
Nonwhite	–0.117 (0.001)	–0.011 (0.003)	–0.03 (0.003)	0.018 (0.005)	Yes	–0.038 (0.002)	–0.068 (0.003)
Pct in occup. with BA	0.508 (0.002)	0.095 (0.005)	0.103 (0.006)	0.007 (0.009)	Yes	0.03 (0.005)	0.011 (0.007)
Constant	2.161 (0.003)	0	0	0		0	0
Observations	2,914,329	2,914,329	1,928,911	1,071,431		2,102,498	1,117,080
R^2	0.19	0.19	0.21	0.13		0.2	0.21

Note: Robust standard errors in parenthesis.

points in wages. Wage growth continues, albeit at a slower clip, throughout one's life.

In the second column, we show the interactions of years of the independent variables with residing in a metropolitan area. We do not report the overall experience coefficients to save space, though they remain similar to those shown in the first column. The coefficients in the second column reflect the extra gains in wages that seem to accrue to metropolitan workers at each experience level. Metropolitan area workers earn a level effect of 0.036 log points more than nonmetro workers at the start of their careers. This gap rises an additional 0.028 log points for workers with between 6 and 10 years of experience. Workers with between 11 and 15 years of experience earn 0.06 log points on top of the level effect, meaning a total premium of 0.096 log points. The coefficients then level off.

These results that replicate those found in Glaeser and Mare (2001) for the 2000 Census, suggest that human capital accumulation is faster in metropolitan areas. The metropolitan area wage effect for inexperienced workers is about one-third its value for more experienced workers. This finding hints at the possibility that much of the effect of cities comes over time, as workers either acquire skills more quickly, or perhaps match more efficiently in large places.

While there is a significant interaction between metropolitan area status and experience, there is no clear link between metropolitan area population and log of experience. In regression (5) of Table 1, we report the absence of such a connection. Workers in a metropolitan area face a steeper age-earnings profile, but the wage profile does not become particularly steep in larger metropolitan areas. In regression (6) of Table 1, we show that being in a skilled area does steepen the age-earnings profile, which is compatible with the view that people are learning more in skilled areas. Regression (7) shows that there is no interaction between metropolitan area population and share with college degrees, which perhaps is unsurprising because there was no experience effect of metropolitan area population.

Returning to Table 3, where there is a basic metropolitan area effect, we now look to see whether there is an interaction between that effect and the skill level of the metropolitan area. Column 3 shows the comparison between those in the 100 most skilled MSAs and those living outside metropolitan areas. Working in these areas provides a large-level effect of 0.069 log points to workers immediately upon starting employment. The experience profile is steeper in these skilled cities than in the full sample shown in column 2. Workers with 6 to 10 years of experience earn an additional 0.035 log point premium, and this rises to 0.075 for those with 11 to 15 years and 0.093 log points for those with 16 to 20 years. These results mean that experience is associated with a 0.162 log point premium for workers in skilled metropolitan areas relative to nonmetropolitan workers.

Column 4 shows that the same does not hold in the 100 least skilled MSAs. Here the level effect is small and insignificant, and the experience trajectory

substantially flatter, showing no significant difference with the nonmetropolitan workers until the 16 to 20 year group. The wage growth associated with living in a metro area comes primarily from highly skilled cities. The *F*-tests in columns 3 and 4 show that the differences in coefficients are significant at the one percent level.

As in our results in Table 1, the fifth and sixth columns find less support for an interaction between city size and experience. Column 5 compares those living in the 25 most populated metropolitan areas to those living in all other metro areas, and finds a significant level effect of 0.081 log points, but no effect on the experience profile. The presence of an interaction between city size and city skill would imply that we might see a stronger effect if we limit the sample to only those in highly skilled cities, but this turns out not to be the case, as Column 6 shows a similar level effect, and no effect on the experience-earnings path.

The results in Table 3 also cast doubt on the view that skilled people are drawn by amenities to locate in larger or more skilled metropolitan areas. If that hypothesis were correct, then the presence of skilled people would act as something of a labor supply shock and we should expect lower earnings for more skilled people in large agglomerations. If amenities were higher in big cities for skilled workers, then the logic of a spatial equilibrium suggests that wages should be lower. Yet across all of our specifications, the interactions between skills and metropolitan locations are positive. As such, it does not seem likely that amenities are causing more skilled people to locate in large metropolitan areas.

More direct evidence on knowledge fails to provide much support to the learning-in-cities hypothesis. In Table 4, we look at the connection between tests of reasoning and vocabulary and both being raised in and currently residing in a city. This evidence is from the General Social Survey (GSS) that subjects adults to tests and has a question about the place in which the adult was brought up. Using place of childhood residence is presumably slightly more exogenous than using place of current residence, but we use both.

The first two columns show that while rural children do worse, the highest test scores were earned by people who were brought up in suburbs. People brought up in big cities do slightly worse on these tests than people brought up in small towns. In the second two columns, we look at place of current residence and find little evidence of a connection between city residence and these skills. While these tests will not capture the most important skills learned working in a big metropolitan area, the fact that we do not find any significant link is not supportive of the learning-in-cities hypothesis.

These results are meant primarily to illustrate the type of evidence that could definitively show that people in cities learn more quickly. So far, no such evidence exists. It is true that people in cities enjoy faster wage growth, but that wage growth is concentrated in more skilled areas. There is no direct evidence linking measurable skill accumulation to urban residence.

TABLE 4: Vocabulary and Reasoning across Places of Residence

	No. Correct on Vocab Test (1)	No. Correct on Reasoning Test (2)	No. Correct on Vocab Test (3)	No. Correct on Reasoning Test (4)
Residence at age 16:				
Rural, nonfarm	-0.261 (0.046)	-0.323 (0.128)		
Rural, farm	-0.423 (0.041)	-0.266 (0.117)		
Small town (under 50,000)	(Omitted)	(Omitted)		
Small city (50,000 - 250,000)	0.024 (0.042)	-0.11 (0.105)		
Suburb of large city	0.283 (0.046)	0.081 (0.108)		
Large city (250,000+)	-0.075 (0.043)	-0.261 (0.116)		
Current residence:				
Rural			-0.108 (0.048)	0.002 (0.152)
Small town (under 50,000)			(Omitted)	(Omitted)
Suburb of small city			0.112 (0.046)	0.177 (0.131)
Small city (50,000-250,000)			-0.134 (0.051)	0.309 (0.135)
Suburb of large city			0.196 (0.044)	0.011 (0.125)
Large city (250,000+)			-0.094 (0.050)	0.147 (0.139)
Years of Schooling				0.203 (0.013)
Age	0.054 (0.004)	0.024 (0.013)	0.05 (0.004)	0.023 (0.013)
Age squared	0 (0.000)	0 (0.000)	0 (0.000)	0 (0.000)
Male dummy	-0.191 (0.027)	0.023 (0.071)	-0.206 (0.027)	0.017 (0.071)
Constant	0.075 (0.114)	0.298 (0.344)	-0.016 (0.115)	0.017 (0.362)
Observations	22,929	2,182	22,970	2,185
R^2	0.27	0.15	0.27	0.14

Notes: (i) Robust standard errors in parenthesis.

(ii) All data for General Social Survey as described in Appendix.

5. CONCLUSION

In this paper, we document that agglomeration effects are much stronger for cities with more skills. This finding points to agglomeration theories that emphasize knowledge accumulation in big cities, rather than theories that emphasize natural advantage or gains from speedy movement of basic commodities. Yet, there is little direct evidence on the knowledge-based agglomeration economies. Empirical researchers have not managed as of yet, to sort out how these agglomeration economies work.

Glaeser and Mare (2001) put forward some evidence suggesting that skill accumulation works faster in metropolitan areas. We duplicate that evidence here, and find that these learning effects are strongest in more skilled metropolitan areas. While these results suggest a strong complementarity between skills, city size, and learning, other direct tests of that complementarity find little evidence. At present we are left with tantalizing hints, but little that is conclusive.

One speculative interpretation of the results is that two things are simultaneously happening in skilled, big cities. First, workers are indeed learning from one another more quickly. Second, the rate of technological change is faster. Together, both effects create the interaction between city size and population across skilled metropolitan areas. The results on age-earnings profiles would be ambiguous if both effects were present, because sometimes the learning effect (which steepens the profile) dominates and sometimes the technological change effect (which flattens the profile) dominates. We hope that further research will sort out these interpretations.

REFERENCES

- Bacolon, M., B.S. Blum, and W.C. Strange. 2010. "Elements of Skill: Traits, Intelligences, and Agglomeration," *Journal of Regional Science*, 50(1), 245–280.
- Berry, C.R. and E.L. Glaeser. 2005. "The Divergence of Human Capital Levels Across Cities," *Papers in Regional Science*, 84(3), August, 407–444.
- Breau, S. and D. Rigby. 2008. "Participation in Export Markets and Productivity of Plants in Los Angeles, 1987–1997," *Economic Geography*, 84(1), January, 27–50.
- Ciccone, A. and R. E. Hall. 2004 [1996]. "Productivity and the Density of Economic Activity," in Paul C. Cheshire and Gilles Duranton (eds.), *Recent Developments in Urban and Regional Economics*, Elgar Reference Collection. International Library of Critical Writings in Economics, Vol. 182, Cheltenham, U.K. and Northampton, MA: Elgar.
- Combes, P.-P., G. Duranton, and L. Gobillon. 2008. "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics*, 63(2), March, 723–742.
- Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux. 2008. "Estimating Agglomeration Economies with History, Geology, and Worker Effects," Ph.D. dissertation. CEPR Discussion Papers.
- Duranton, G. and M.A. Turner. 2008. "Urban Growth and Transportation," Ph.D. dissertation. CEPR Discussion Papers.
- Ellison, G. and E.L. Glaeser. 1999. "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?" *American Economic Review*, 89(2), May, 311–316.
- Freedman, M.L. 2008. "Job Hopping, Earnings Dynamics, and Industrial Agglomeration in the Software Publishing Industry," *Journal of Urban Economics*, 64(3), November, 590–600.

- Garofalo, G.A. and S. Yamarik. 2002. "Regional Convergence: Evidence from a New State-by-State Capital Stock Series," *Review of Economics and Statistics*, 84(2), (May, 316–323.
- Glaeser, E.L. 1999. "Learning in Cities," *Journal of Urban Economics*, 46(2), September, 254–277.
- . 2005. "Urban Colossus: Why is New York America's Largest City?" *Federal Reserve Bank of New York Economic Policy Review*, 11(2), December, 7–24.
- . 2008. *Cities, Agglomeration and Spatial Equilibrium*, Oxford: Oxford University Press.
- Glaeser, E.L. and J.D. Gottlieb. 2008. "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, Spring 2008, 155–239.
- Glaeser, E.L. and D.C. Mare. 2001. "Cities and Skills," *Journal of Labor Economics*, 19(2), April, 316–342.
- Glaeser, E.L. and A. Saiz. 2004. "The Rise of the Skilled City," *Brookings-Wharton Papers on Urban Affairs*, 2004, 47–94.
- Kolko, J. and D. Neumark. 2008. "Changes in the Location of Employment and Ownership: Evidence from California," *Journal of Regional Science*, 48(4), October, 717–743.
- Krugman, P. 1991. "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), June, 483–499.
- Lucas, R.E., Jr. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), July, 3–42.
- Marshall, A. 1890. *Principles of Economics*, London: Macmillan.
- Moretti, E. 2008. "Real Wage Inequality," Ph.D. dissertation, NBER Working Papers. National Bureau of Economic Research, Inc.
- Munnell, A.H. 1991. "Is There a Shortfall in Public Capital Investment? An Overview," *New England Economic Review*, May–June, 1991, 23–35.
- Peri, G. 2002. "Young Workers, Learning, and Agglomerations," *Journal of Urban Economics*, 52(3), November, 582–607.
- Puga, D. 2010. "The Magnitude and Causes of Agglomeration Economies," *Journal of Regional Science*, 50(1), 203–219.
- Shapiro, J.M. 2006. "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital," *Review of Economics and Statistics*, 88(2), May, 324–335.

APPENDIX: DATA

For Figures 1, 3 and 4, and Equations (1) and (2), productivity (or output per worker) is calculated by dividing the Gross Metropolitan Product for 2001 (from the Bureau of Economic Analysis at <http://www.bea.gov/regional/gdpmetro/>) by the total labor force for 2000 (from published 2000 Census figures). Population and share with BAs also comes from the published 2000 Census figures, and this population and BA data are also used in Figure 2. For Figure 3, "less skilled Metropolitan Statistical Areas (MSAs)" refer to those MSAs that have the share of the population with BAs in 2000 less than 17.64 percent. For Figure 3, "more skilled MSAs" refer to those MSAs that have the share of the population with BAs in 2000 more than 25.025 percent.

For (2'), population and share with BAs in 1940 comes from published 1940 Census figures. For (2''), population in 1900 comes from published 1900 Census figures. For (3), real family income is calculated using family median income from the published 2000 Census figures, divided by the cost-of-living index for each MSA published by the American Chamber of Commerce Research Association at <http://www.coli.org/>. Data for Figure 5 are calculated using the 2006 *Annual Survey of Manufactures*, with details described in the paragraph about Table 2 below.

The individual-level data used in Tables 1 and 3 come from the Integrated Public Use Microdata Series (IPUMS) 2000, five percent Census sample. Where aggregate metro-area numbers such as population and the percentage of workers over 25 with a college degree are used in conjunction with individual-level data, these are merged on from published Census figures, since the IPUMS does not fully identify all metro areas. All individual-level regressions are run for male workers aged 25 to 65. Hourly earnings are calculated by dividing yearly earned income by number of weeks worked and usual weekly hours. Experience is calculated as age minus years of schooling minus six, where years of schooling is approximated as precisely as possible using the categorical schooling variable provided in the 2000 Census. All calculations are weighted by person weight unless otherwise noted.

In Table 2, the state-level log capital per worker, log value added per worker and log hourly wage were calculated using the total capital expenditures, total value added, number of employees, total production workers wages, and total production workers hours data from the 2006 *Annual Survey of Manufactures* at factfinder.census.gov. The state-level density and years of schooling variables come directly from Table 2 of "Productivity and the Density of Economic Activity" by Antonio Ciccone and Robert E. Hall, *American Economic Review*, March 1996.

Data for Table 4 come from the General Social Survey. In 1994, eight reasoning questions were asked that required the respondent to assess the similarities between various objects and ideas. Their responses were coded as correct, partly correct, or incorrect by the GSS, and information on this coding is available in Appendix D of the GSS cumulative codebook. Our dependent variable is the number of fully correct responses out of the eight questions. The vocabulary test, given in 17 waves of the survey spaced between 1974 and 2006, asks the respondent 10 vocabulary questions and records the number of correct responses. We pool across the waves, weighting using the WTSSALL variable which, which adjusts for the ways that the counting of households in the survey has changed over time. For the residence variables, categories of *xnorsize* (current residence) are combined so as to mirror the categories of the *res16* (residence at age 16) variable as closely as possible.